典型的全连接层神经网络的输入和输出用x = [x1,x2，…]表示。
，X100]7和y = [y1,y2，…，y98]7。
在该层中应用线性激活函数，连接输入和输出的网络参数由大小为98×100的矩阵W =[Wij]和98维向量θ =[θ1，θ1，…]
，θ98]T。

3. The inputs and outputs of a typical fully-connected layer of neural network are denoted by $\boldsymbol{x} = [x_1, x_2, \ldots, x_{100}]^T$ and $\boldsymbol{y} = [y_1, y_2, \ldots, y_{98}]^T$. A linear activation function is applied in this layer and the network parameters to connect the inputs and outputs are given by a matrix $\boldsymbol{W} = [w_{ij}]$ of size 98×100 and a 98-dimensional vector $\boldsymbol{\theta} = [\theta_1, \theta_1, \ldots, \theta_{98}]^T$.

(a) (i) Express the outputs in term of inputs in vector-matrix form and scalar form.
用向量矩阵形式和标量形式表示输入的输出。

(ii) Compute the numbers of trainable parameters, multiplications and summations required in this layer to compute the outputs from the inputs.
计算可训练参数的数量，这一层需要的乘法和求和来计算输入的输出。

(iii) What is the ratio of the number of outputs to the number of trainable parameters?
输出的数量与可训练参数的数量之比是多少？

(12 Marks)

将该层替换为卷积神经网络层，该层有20个可学习的过滤器，大小为3，
具有可训练的参数，[]
这将生成20个输出特征映射，用[]表示。
在这一层中应用线性激活函数。

(b) Replace this layer by a convolutional neural network layer that has 20 learnable filters of size 3 with trainable parameters, $\boldsymbol{w}^k = [w_{-1}^k, w_0^k, w_1^k]^T$ and θ^k, $1 \le k \le 20$. This generates 20 output feature maps, denoted by $\boldsymbol{y}^k = [y_2^k, y_3^k, \ldots, y_{99}^k]^T$, $1 \le k \le 20$. A linear activation function is applied in this layer.

(i) Express the outputs in term of inputs in scalar form. 用标量形式表示输入的输出。

(ii) Compute the number of trainable parameters, multiplications and summations required in this layer to compute the outputs from the inputs.
计算可训练参数的数量，这一层需要的乘法和求和来计算输入的输出。

(iii) What is the ratio of the number of outputs to the number of trainable parameters?
输出的数量与可训练参数的数量之比是多少？

(13 Marks)

23-S1-Q3

(a)(i) Q: vector-matrix form and scalar form input and output?

$x = [x_1\ x_2\ x_3 \cdots x_{100}]^T$

$y = [y_1\ y_2 \cdots y_{98}]^T$

$W = [w_{ij}]$ 98×100

$\theta = [\theta_1\ \theta_2 \cdots \theta_{98}]^T$

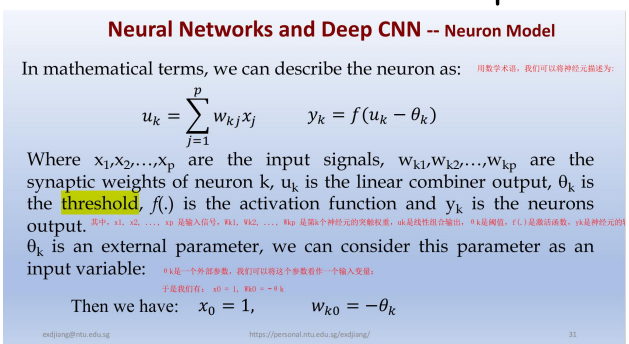
vector-matrix form

$y = Wx + \theta$

scalar form

$y_i = \sum_{j=1}^{100} w_{ij} x_j + \theta_i \quad i = 1, 2, \cdots 98$

$y_1 = w_{11}x_1 + w_{12}x_2 + \cdots + \theta_1$



第一题是+θ还是-θ？

(a) Output after FC layer

$: Wx + b$

$: \begin{bmatrix} 0 & 3 & 7 & 8 \\ 1 & 8 & 0 & 0 \\ 0 & 8 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.3 \\ 0 \\ 0.2 \\ 0.1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.2 \\ 1.3 \\ 2.2 \end{bmatrix}$

(ii) Q parameters? multiplications? summations?

Solution

parameters: $98 \times 100 + 98 = 9898$

multiplications: $98 \times 100 = 9800$

summations: $98 \times (99+1) = 9800$

100数加99次 加1次θ

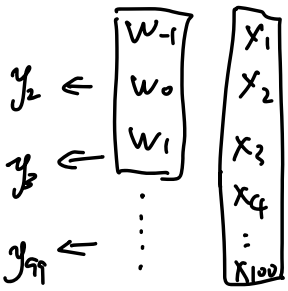
(iii) Q: ratio

Solution Ratio $= \frac{98}{9898} \approx 0.0099$

(b)(i) $y_i^{(k)} = w_{-1}^{(k)} x_{i-1} + w_0^{(k)} x_i + w_1^{(k)} x_{i+1} + \theta^{(k)}$

$k = 1, 2, \cdots, 20$

$i = 1, 2, \cdots 99$



(ii) parameters: $20 \times (3+1) = 80$

20个(3权重+1个bias)

multiplications: $3 \times 98 \times 20 = 5880$

每个y乘3次，98个y共20个

summations: $(2+1) \times 98 \times 20 = 5880$

每个y加2权加1θ

98个y，20个

(iii) Ratio $= \dfrac{98 \times 20}{80} = 24.5$